Common Language for Assessment in Utah

Assessment - An educational practice with the purpose of collecting evidence of student learning.

Authentic Assessment – Assessments that require students to perform complex tasks representative of activities actually done in out-of-school settings (NRC, 2001). These assessments require students to apply or perform complex tasks.

Constructed Response – Item types based on a prompt (eg. directions, questions, and problems) that require students to generate a response (eg. written and pictorial) on their own. Sometimes called an "open-ended" item (CTB-McGraw Hill, 2003).

Content Standard – A broad statement of what students are expected to know and be able to do (USOE, 2002).

Core/Core Curriculum – Those standards of learning, as well as the ideas, concepts, and skills that provide a foundation on which subsequent learning may be built, that are essential for all Utah students, (USOE, 2002). This curriculum includes Intended Learning Outcomes (ILOs), Benchmarks, Standards, and Objectives.

CRT – Criterion-Referenced Test – A test to measure performance against a specific standard (USOE, 2002).

Diagnostic Test - A test intended to locate learning difficulties or patterns of error. Such tests yield measures of specific knowledge, skills, or abilities underlying achievement within a broad subject. Thus, they provide a basis for remedial instruction (CTB-McGraw Hill, 2003).

Evaluation - A process of <u>comparing</u> student's performance or characteristics against a standard. Assessment followed by making value judgments about pupils' performance. Students are compared with other students or with a predetermined standard (Baird, 1997).

Formative Assessment - Assessment which provides learning experiences for students as well as feedback for teachers. Assessment which measures student learning in order to identify how well they are learning or how much of the subject matter they have mastered in order to help them learn more or to help the teacher to improve ongoing instruction (Baird, 1997). All those activities undertaken by teachers and by their students that provide information to be used as feedback to modify the teaching and learning activities in which they are engaged (Black and William, 1998 from NRC, 2001).

Grade – A symbol representing a level of performance; this is usually based on a vague standard.

Inference – A conclusion drawn based on data.

Interpretive Exercise – Item type that consists of a set of data or information, a display, followed by a series of problems or questions having answers that are dependent upon the information given. These items are also called text dependent (Baird, 1997).

Item – A question or set of instructions that require a student response.

Measurement – A process of <u>collecting</u> information about the performance of a student or a class. Measurement often includes the assignment of a number to express in quantitative terms the degree to which a pupil possesses a given characteristic (Baird, 1997).

Multiple Choice – Item type that consists of a problem and a list of suggested solutions (Baird, 1997).

NRT – Norm Referenced Test – A test where scores are based on comparisons with a nationally representative group of students (USOE, 2002).

Performance Based Assessment – Assessments that allow students to demonstrate their understandings and skills as they perform a certain activity (NRC, 2001). Usually there are multiple ways that an examinee can approach a performance assessment and more than one correct answer (CTB-McGraw Hill, 2003).

Performance Standards – Points on a test score scale that are used in separating examinees into performance categories (Cizek, 2001).

Performance Task –A performance based assessment which requires at least two different sessions in class to complete.

Performance Test – A performance based assessment that can be completed in one session of class.

Reliability – The consistency of test scores obtained by the same individuals on different occasions or with different sets of equivalent items; accuracy of scores (CTB- McGraw Hill, 2003).

Rubric – A scoring tool, or set of criteria, used to evaluate a student's performance (CTB-McGraw Hill, 2003) on a test item or test.

Selected Response – Item types that have students choose from provided options. **Summative Assessment** - Evaluation which tests students' performance to determine students' final overall assimilation of course material and/or overall instructional effectiveness (Baird, 1997).

Test – Measurement instrument designed to obtain a sample of an individuals' optimal performance in a specified content area (Crocker and Algina, 1986).

Validity – A term referring to a test measuring what is purports to measure and allowing scores to be used to draw appropriate inferences (CTB-McGraw Hill, 2003). The aspects of validity include: content, criterion, construct, and consequential.

References

Baird, H. (1997). Performance Assessment for Science Teachers. SLC, UT: USOE.

Cizek, Gregory J. (2001). Setting Performance Standards: Concepts, Methods, and Perspectives. Mahwah. NJ: Lawrence Erlbaum Associates.

Crocker, L. & Algina, J. (1986). *Introduction to Classical & Modern Test Theory*. Belmont, CA: Wadsworth Thomson Learning.

CTB-McGraw Hill, (2003). Glossary of Assessment Terms [On-line]. Available:

http://www.ctb.com/articles/article_information.jsp?CONTENT%3C%3Ecnt_id=38937&FOLDER%3C%3Efolder_id=39863&bmUID=1074276256685

National Research Council (NRC). (2001). *Classroom Assessment and the National Science Education Standards*. J.M. Atkin, P. Black, and J. Coffey (Eds.), Committee on Classroom Assessment and the National Science Education Standards. Washington, DC: National Academy Press.

USOE, (2002). Secondary Core Curriculum. SLC, UT, Utah State Office of Education.